

Perbandingan Metode ARIMA (*Box Jenkins*) dan Metode *Winter* dalam Peramalan Jumlah Kasus Demam Berdarah Dengue

Metta Octora dan Kuntoro

Departemen Biostatistika dan Kependudukan FKM UNAIR

Fakultas Kesehatan Masyarakat Universitas Airlangga

Alamat korespondensi:

Kuntoro

E-mail: kuntoro2@yahoo.com

Departemen Biostatistika dan Kependudukan FKM UNAIR

Fakultas Kesehatan Masyarakat Universitas Airlangga

Kampus C Unair Mulyorejo Surabaya 60115

ABSTRACT

A good forecasting method is a method that has the smallest error rate in forecasting. Each ARIMA (Box Jenkins) and Winter method have advantages and disadvantages when compared with other methods. For comparing these methods, we used Dengue Haemorrhagic Fever (DHF) case because of seasonal feature. The study has been done to compare ARIMA and Winter method by determining the best mathematical model, and the smallest prediction error on the number of DHF cases in Surabaya. The data was DHF case at Health Department of Surabaya for the period from January 2005 until June 2010. Time series data are classified monthly that are known have cyclic periodic movements. Earlier variants should be tested first by comparing the individual values with the average value for each year. If the data is already seasonal then analyzed with Winters and ARIMA method. Winters method used 4 models, while ARIMA method obtained 3 models. Furthermore, mathematical models are determined the smallest forecasting error rate by the smallest value MAPE, MAD and MSE indicator to predict the incidence of DHF in the next 6 months. The smallest error in sample value of Winters method is model 3 with MAPE 49.14212; MAD 88.5205; and MSE 18322.02, while the smallest error out sample value of Winters method is a model 4 with MAPE 3.8810; MAD 17.4669 and MSD 4535.979982. The smallest error in sample value of ARIMA method is model 1 with MAPE 3.9667, MAD 0.1935 and MSD 0.067899. The smallest error out sample value of ARIMA method is model 2 with MAPE 1.0286; MAD 0.0620 and MSD 0.0489032 of the these methods are analyzed can be concluded that the method of ARIMA (1,0,2) (1,0,2) is the best method because it has the MAPE, MAD and MSD is smaller than the method of Winter with parameters $\alpha = 0.2$, $\gamma = 0.15$ and $\delta = 0.002$.

Keywords: ARIMA, winters, seasonal, MAPE, MAD, MSD

ABSTRAK

Metode peramalan yang baik adalah metode yang mempunyai tingkat kesalahan terkecil dalam peramalan. Metode ARIMA dan *Winter* masing-masing mempunyai kelebihan dan kekurangan bila dibandingkan dengan metode lainnya. Dalam membandingkan kedua metode, yakni ARIMA dan *winter*, digunakan kasus Demam Berdarah Dengue (DBD) karena kasus ini bersifat musiman. Tujuan penelitian adalah membandingkan metode *Box Jenkins* (ARIMA) dengan metode *Winter* dengan menentukan model matematis terbaik, serta menentukan kesalahan ramalan terkecil pada jumlah kasus DBD di Dinas Kesehatan Kota Surabaya. Data yang digunakan adalah kasus DBD di Dinas Kesehatan Kota Surabaya untuk periode Januari 2005 sampai dengan Desember 2009. Data time series diklasifikasikan ke dalam periode bulanan sehingga diketahui mempunyai gerakan periodik yang berulang. Sebelumnya harus diuji terlebih dahulu variasinya dengan membandingkan nilai individual dengan nilai rata-rata untuk tiap tahunnya. Bila data sudah bersifat *seasonal* maka dilakukan analisis dengan Metode *Winters* maupun Metode *Box Jenkins*. Pada metode *Winters* dihasilkan 4 model, sedangkan pada metode ARIMA diperoleh 3 model. Selanjutnya ditentukan model matematis dengan tingkat kesalahan peramalan terkecil yaitu dengan indikator MAPE, MAD maupun MSD terkecil untuk meramalkan kejadian DBD dalam 6 bulan mendatang. Nilai galat error data *insample* Metode *Winters* terkecil adalah model 3 dengan MAPE 49,14212, MAD 88,5205 dan MSE 18322,02; sedangkan nilai galat error data *outsample* Metode *Winters* terkecil adalah model 4 dengan MAPE 3,8810, MAD 17,466 9 dan MSE 4535,979982. Nilai galat error data *insample* metode ARIMA terkecil adalah model 3 dengan MAPE 3,9667, MAD 0,1935 dan MSD 0,067899. Nilai galat error data *outsample* metode ARIMA terkecil adalah model 2 dengan MAPE 1,0286, MAD 0,0620 dan MSD 0,0489032. Dari kedua metode yang dianalisis dapat

disimpulkan bahwa Metode ARIMA (2,0,1)(2,0,0) merupakan metode terbaik karena mempunyai MAPE, MAD dan MSD yang lebih kecil dibandingkan dengan Metode *Winter* dengan parameter $\alpha = 0,2$, $\gamma = 0,15$ dan $\delta = 0,002$

Kata kunci: ARIMA, *winter*, musiman, MAPE, MAD, MSD

PENDAHULUAN

Peramalan merupakan suatu proses yang bertujuan menduga suatu kejadian yang akan datang dan merupakan alat bantu yang penting dalam perencanaan yang efektif dan efisien (Makridakis *et al.*, 1999). Metode peramalan merupakan metode yang berfungsi untuk memprediksi data runtun waktu beberapa periode yang akan datang berdasarkan data di masa lalu, yang dinyatakan dalam bentuk persamaan matematis.

Metode peramalan melalui analisis suatu variabel yang akan diperkirakan dengan variabel atau fungsi waktu, dikenal sebagai metode hubungan *time series*. Analisis data *time series* dapat digunakan untuk mengetahui perkembangan satu atau beberapa kejadian dan hubungan atau pengaruhnya pada kejadian lainnya, sekaligus menghitung secara kuantitatif besarnya pengaruh suatu kejadian dengan kejadian lainnya apabila berhubungan. Data *time series* juga dapat digunakan untuk membuat ramalan berdasarkan garis regresi atau *trend*.

Menurut Box dan Jenkins (1976), metode ARIMA (*Autoregressive Integrated Moving Average*) atau juga dikenal dengan metode Box-Jenkins merupakan salah satu bentuk analisis deret waktu (*time series analysis*). Model ARIMA dapat menganalisis data secara univariat yang mengandung pola musiman maupun *trend*. Metode ini hanya menganalisis data yang stasioner, sehingga data yang tidak stasioner harus distasionerkan terlebih dahulu dengan transformasi atau pembedaan. Berbeda dengan metode pemulusan eksponensial yang dapat menganalisis berbagai data baik data yang stasioner maupun data yang tidak stasioner dengan pemisahan aspek *trend* dan musiman (Kurniawan, 2008).

Hal utama yang menjadi ciri dari model ARIMA dalam rangkaian analisis data deret waktu dibandingkan metode pemulusan adalah perlunya pemeriksaan keacakan data dengan melihat koefisien autokorelasinya. Model ARIMA

juga bisa digunakan untuk mengatasi masalah sifat keacakan, *trend*, musiman bahkan sifat siklis data *time series* yang dianalisis.

Selain metode ARIMA juga dikenal metode pemulusan *Winter*, yang merupakan metode peramalan *trend* dan musiman. Jadi jika data yang akan dianalisis merupakan data musiman maka diperlukan sebuah metode pemulusan yang dapat menangani faktor musiman secara langsung, yaitu metode *Winters*. Metode *Winters* sebenarnya serupa dengan metode Holt, namun ditambah dengan sebuah persamaan guna mengatasi variasi musiman statistik yang digunakan untuk menggambarkan hubungan antara satu variabel respons (*dependent*) dengan satu atau lebih variabel prediktor (*independent*). Pada dasarnya metode *Winters* menggunakan data masa yang dimuluskan dengan cara melakukan pembobotan menurun secara eksponensial terhadap nilai pengamatan yang lebih tua atau nilai yang lebih baru diberikan bobot yang relatif lebih besar dibanding nilai pengamatan yang lebih lama.

Metode ARIMA dan *Winter* masing-masing mempunyai persamaan yaitu keduanya menganalisis data secara univariat yang mengandung pola musiman dan *trend*. Keduanya juga mengasumsikan nilai dan kesalahan masa lalu sebagai dasar peramalan di masa mendatang.

Tabel 1. Perbedaan Metode ARIMA dan *Winter*

ARIMA	<i>Winter</i>
1. Hanya pada data stasioner	1. Dapat untuk data stasioner dan bukan
2. Perlu uji keacakan data dengan melihat koefisien autokorelasi	2. Tidak perlu
3. Berdasar analisis pemilihan model <i>trend</i> musiman ARIMA	3. Berdasar analisis model regresi deret waktu sederhana

Namun demikian masing-masing metode juga mempunyai perbedaan seperti yang tercantum dalam Tabel 1.

Dalam penelitian ini metode ARIMA ingin dibandingkan dengan metode *Winter* karena setiap metode peramalan memiliki keunggulan dan kelemahan dalam menganalisis data, sehingga diperoleh bentuk model terbaik dengan tingkat kesalahan terkecil dalam meramalkan suatu kasus.

Dalam membandingkan kedua metode, digunakan kasus Demam berdarah *dengue* (DBD) sebagai variabel data. DBD bersifat musiman sehingga cocok bila digunakan untuk membandingkan ARIMA dan *Winters* yang bersifat musiman, selain itu DBD masih merupakan masalah kesehatan masyarakat yang menimbulkan dampak sosial maupun ekonomi. Jumlah kasus yang dilaporkan cenderung meningkat dan daerah penyebarannya bertambah luas (Depkes RI, 2005).

Sejak pertama ditemukan penyakit DBD di Indonesia (Surabaya dan Jakarta) pada tahun 1968, jumlah kasus cenderung meningkat dan daerah penyebarannya bertambah luas, sehingga pada tahun 1994 DBD telah tersebar ke seluruh Propinsi di Indonesia. Pada 1997–2001 jumlah rata-rata kasus dilaporkan sebanyak 40.854 kasus dengan rata-rata kematian 701 orang setiap tahunnya. Pada tahun yang sama setiap 100.000 penduduk, 20–21 orang diantaranya menderita DBD dan setiap 100 penderita, rata-rata yang meninggal sebanyak 1–2 orang (Depkes RI, 2006).

Kejadian Luar Biasa Demam Berdarah *Dengue* (KLB DBD) terbesar terjadi pada tahun 1998, dengan *insiden rate* (IR) = 35,19 per 100.000 penduduk dan Crude Fatality Rate (CFR) 2 %. Pada tahun 1999 IR menurun tajam sebesar 10,17 %, namun pada tahun-tahun berikutnya IR cenderung meningkat yaitu 15,99 % (2000); 21,66 % (2001); 19,24 % (2002); dan 23,87 % (2003) (Depkes RI, 2004).

Berdasarkan masalah tersebut maka ingin dibuat suatu model peramalan yang akurat dengan kesalahan terkecil untuk memprediksi kasus DBD dalam kurun waktu tertentu di masa mendatang sehingga dapat menentukan manajemen yang tepat dan efektif, baik dari segi pendanaan, fasilitas medis maupun tenaga medis,

mengingat kasus ini meningkat secara tajam pada musim tertentu dan menimbulkan dampak morbiditas dan mortalitas.

Tujuan penelitian ini adalah membandingkan metode ARIMA dengan metode *Winter* dalam memprediksi kasus DBD, mengetahui langkah dalam menentukan model matematis serta menentukan kesalahan ramalan terkecil pada kasus DBD sehingga mengetahui akurasi peramalan kedua metode tersebut.

METODE PENELITIAN

Penelitian dilakukan dengan mengambil data sekunder kasus DBD di Surabaya untuk periode Januari 2005 sampai dengan Juni 2010 yang diamati secara runtut (berkala) setiap bulan. Variabel dependen dalam penelitian ini adalah kasus DBD (Y), sedangkan variabel independennya adalah waktu (X).

Analisis data dilakukan secara statistik dengan bantuan program komputer. Tahapan analisis data metode ARIMA meliputi: Eksplorasi data, Tahapan Metode ARIMA, dan membandingkan nilai kesalahan peramalan terkecil antara metode ARIMA dengan *Winters* pada kasus DBD dengan menghitung *MAPE*, *MAD* dan *MSD*.

Tahapan Metode ARIMA sendiri terdiri dari beberapa tahap yaitu: (a) Tahap identifikasi dengan pengidentifikasian model yang dianggap paling sesuai dengan melihat plot ACF dan PACF dari *correlogram*; (b) Tahap estimasi parameter dengan penaksiran terhadap parameter dalam model tersebut; (c) Tahap uji diagnostik: untuk menguji kesesuaian dari parameter yang didapat pada tahap sebelumnya; (d) Tahap peramalan/penerapan: dilakukan setelah model yang sesuai teridentifikasi.

Tahapan Metode *Winters* meliputi: (a) mengidentifikasi model; (b) menentukan nilai awal taksiran parameter; (c) menentukan konstanta pemulusan (α, γ, β); (d) menghitung nilai ramalan data asli; (e) meramalkan periode mendatang.

HASIL PENELITIAN

Tabel 2 menunjukkan data *insample* kasus DBD meningkat tajam pada sekitar bulan Maret,

April dan Mei setiap tahunnya kecuali pada tahun 2005 yang mana terjadi puncak peningkatan terjadi pada bulan Agustus.

Pada deskripsi statistik dari jumlah kasus DBD yang terdapat pada Tabel 2 dan Gambar 6 terlihat bahwa rata-rata jumlah kasus DBD paling tinggi dicapai pada bulan April, yaitu sebesar 409,2 kasus. Pada bulan April kota Surabaya sedang berada pada musim penghujan. Variansi yang terbesar terjadi pada bulan Maret. Jumlah kasus DBD tercatat tertinggi pada tahun 2006 terjadi pada Maret, yaitu sebesar 841 kasus, sedangkan jumlah kasus terkecil yakni sebesar 40 kasus terjadi pada bulan Oktober.

Hasil Metode *Winters*

Dari Gambar 1, terlihat bahwa data memiliki pola musiman dan trend, demikian pula dari deskripsi statistiknya dapat diketahui pola musimannya adalah mengikuti pola musiman 12 bulanan. Dari plot Gambar terlihat bahwa data bersifat trend multiplikatif yang mempunyai kecenderungan menurun mulai tahun 2006 sampai 2009.

Dengan menggunakan Metode *Winter* multiplikatif diperoleh 4 model Dari Gambar 2 dapat dijelaskan bahwa model 1 metode

Winters yang menggunakan parameter uji coba $\alpha = 0,2$; $\gamma = 0,1$; $\delta = 0,005$ diperoleh nilai FITS yang merupakan nilai peramalan. Nilai FITS selanjutnya dimasukkan ke dalam rumusan galat error yang meliputi MAPE, MAD, MSE. *MAPE* secara umum dirumuskan sebagai berikut:

$$MAPE = \frac{\sum |(y_t - \hat{y}_t)/y_t|}{n} \times 100$$

Rumus MAD adalah sebagai berikut:

$$MAD = \frac{\sum |y_t - \hat{y}_t|}{n}$$

Rumus MSD adalah sebagai berikut:

$$MSD = \frac{\sum (y_t - \hat{y}_t)^2}{n}$$

Di mana:

y_t = nilai data sesungguhnya pada periode ke-t

$\hat{y}_t$ = nilai ramalan pada periode ke-t

n = banyaknya data

Dari parameter model 1 ($\alpha = 0,2$; $\gamma = 0,1$; $\delta = 0,005$); model 2 ($\alpha = 0,2$; $\gamma = 0,15$; $\delta = 0,002$); model 3 ($\alpha = 0,2$; $\gamma = 0,08$; $\delta = 0,005$); model 4 ($\alpha = 0,2$; $\gamma = 0,15$; $\delta = 0,002$) diperoleh hasil *insample* MAPE, MAD dan MSD seperti yang diperlihatkan Tabel 2. Data aktual *outsample* kasus demam berdarah (periode 61–66) mulai Januari 2010 sampai dengan Juni 2010 digambarkan dengan diagram plot time series seperti yang terlihat pada Gambar 1 menunjukkan kenaikan *peak*.

Perhitungan galat error data *outsample* model 1, 2, 3, dan 4 dilakukan dengan memasukkan nilai FORECAST (yang juga merupakan nilai peramalan), bersama dengan nilai aktual *outsample* ke dalam rumus MAPE, MAD dan MSD sehingga diperoleh nilai galat error data *outsample*. Dari perhitungan nilai galat error data *outsample* Metode *Winters* model 1, 2, 3, dan 4 diperoleh nilai MAPE, MAD dan MSD seperti yang diperlihatkan pada Tabel 2.

Dari keempat model *insample* kasus demam berdarah dengue yang menggunakan metode *Winters* dapat disimpulkan bahwa model yang terbaik adalah model 3, yaitu model yang

Tabel 2. Deskripsi Statistik DBD di Surabaya

Bulan	Mean	SD	Min	Max
1	240,2	188,26	74	557
2	377,8	246,99	123	790
3	391,2	259,53	183	841
4	409,2	184,53	163	672
5	383,8	183,35	169	624
6	278	106,52	165	441
7	190	61,95	109	261
8	206	176,89	66	505
9	132,6	125,00	54	352
10	77,4	59,26	40	182
11	90,2	61,52	39	189
12	104,8	102,66	29	285
Mean keseluruhan			240,1	
SD keseluruhan			189,5737145	
Jumlah kasus terkecil			74	
Jumlah kasus terbesar			8	

memiliki MAPE, MAD dan MSD terkecil dibandingkan dengan ketiga model data *insample* lainnya. Dari Tabel 2 juga disimpulkan bahwa dari keempat model data *outsample* yang digunakan, model yang terbaik adalah model 4 karena mempunyai MAPE, MAD dan MSD yang terkecil bila dibandingkan dengan ketiga model data *outsample* lainnya.

Namun, dari Gambar 1 juga terlihat bahwa nilai jumlah kasus terhadap standar deviasinya mempunyai tinggi rendahnya tidak sama untuk setiap bulannya. Hal ini menunjukkan bahwa data tidak stasioner di dalam varians. oleh karena data stasioner di dalam mean namun tidak stasioner di dalam varians maka perlu dilakukan transformasi.

Transformasi Data Time Series

Data *time series* baik *insample* maupun *outsample* yang belum stasioner dalam varians tersebut selanjutnya ditransformasi menggunakan transformasi *Box Cox* dengan bantuan software komputer. Hasil transformasi ditunjukkan dalam plot time series (Gambar 2).

Untuk melakukan uji hipotesis pada kasus ini sebelum dan sesudah ditransformasi digunakan *Bartlett's test* dan *Levene's test*, kemudian dicari hasil yang terbaik dari kedua uji tersebut. Uji normalitas diperlukan sebelum dilakukan *Bartlett's test*. Dari Gambaran uji normalitas diperoleh bahwa hasil uji normalitas sebelum transformasi mempunyai nilai $p < 0,000$ yang berarti bahwa data tidak terdistribusi normal berarti tidak dapat menggunakan *Bartlett's test*, jadi menggunakan acuan *Levene's test*. Sedangkan pada uji normalitas setelah transformasi diperoleh nilai $p > 0,150$, berarti *Bartlett's test* dapat digunakan.

Dari hasil *Bartlett's test* setelah transformasi diperoleh nilai p adalah 0,124. Karena nilai p lebih besar dari 0,05 maka berarti menerima hipotesis nol. Jadi disimpulkan bahwa data setelah ditransformasi homogen.

Secara garis besar dari uji di atas diperoleh hasil bahwa transformasi *Box Cox* berhasil menstasionerkan varians dengan nilai lambda 0,19 (95% CI). Nilai lambda ini selanjutnya akan digunakan pada saat mengembalikan bentuk data transformasi menjadi bentuk data semula pada akhir tahap peramalan.

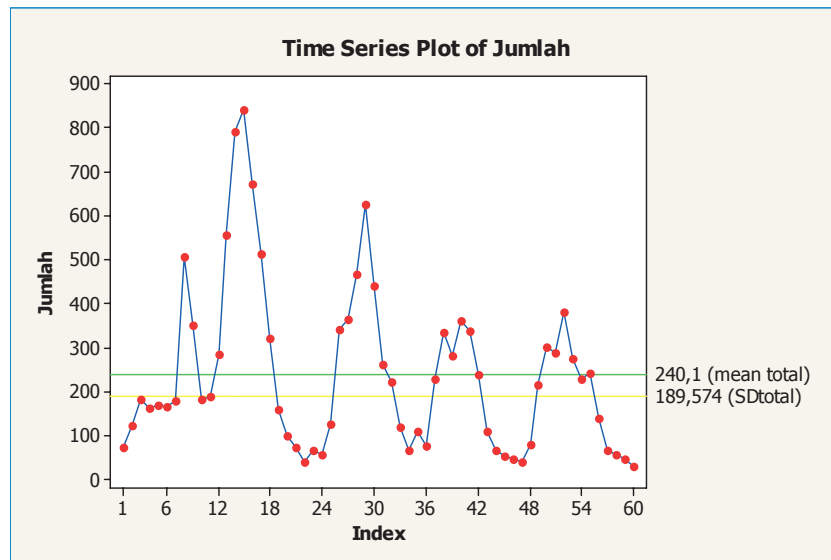
Pada tahap identifikasi selanjutnya ditentukan ACF dan PACF. Dari hasil ACF data transformasi (Gambar 3) diperlihatkan bahwa AR signifikan pada lag 1, lag 2, lag 6 dan lag 7, karena memotong garis signifikansi. Tampak pula bahwa plot data tersebut prosesnya turun secara cepat mendekati nol pada lag 2, hal tersebut menunjukkan stasioneritas. Dari Gambaran PACF (Gambar 4) diperoleh hasil bahwa MA signifikan pada lag 1, lag 2 dan lag 4, karena memotong garis signifikansi. Pada PACF terjadi penurunan secara cepat mendekati nol pada lag 1, hal tersebut juga menunjukkan stasioneritas.

Setelah data dinyatakan telah stasioner maka dilakukan estimasi parameter sementara dengan berdasarkan plot ACF dan PACF. Dari plot ACF terdapat empat koefisien yang signifikan (memotong garis signifikansi), sehingga dapat disimpulkan sementara nilai untuk AR adalah 1, 2, 6 dan 7. Sedangkan pada plot PACF terdapat koefisien yang signifikan (memotong garis silang kepercayaan), yaitu 1, 2 dan 4. Pada kasus ini tidak dilakukan *differencing*, karena sudah dilakukan transformasi pada data awal. Dari kedua fungsi tersebut dapat diestimasi beberapa model ARIMA yaitu model 1 (2,0,2)(2,0,2); model 2 (1,0,2)(1,0,2) dan model 3 (1,0,1)(1,0,1) dengan periode 12 bulan. Pada Tabel 2 diperlihatkan hasil output data *insample* dan data *outsample* yang menggunakan Metode ARIMA.

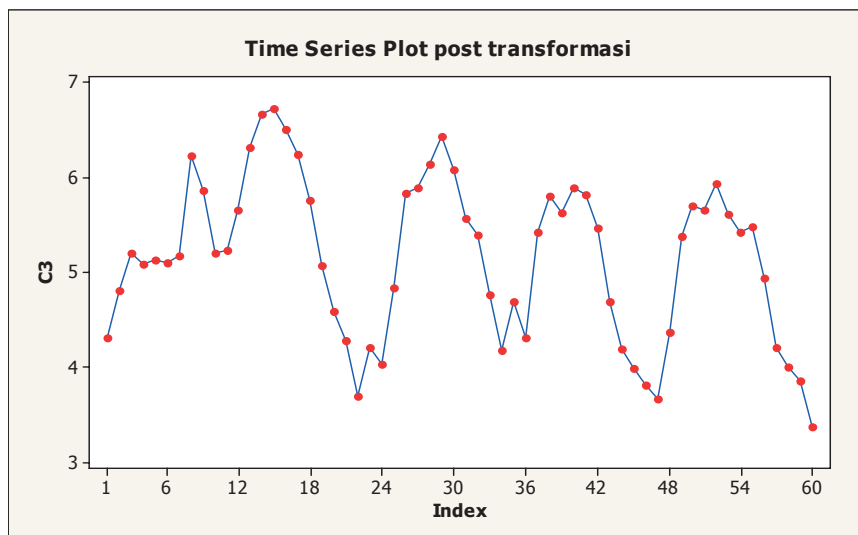
Pada Tabel 3 diperlihatkan bahwa dari ketiga model *insample* yang diperoleh model yang terbaik adalah model 1, yaitu ARIMA (2,0,2)(2,0,2). Dari ketiga model data *outsample* diperoleh hasil bahwa model yang terbaik adalah model 2, yaitu ARIMA (1,0,2)(1,0,2). Selanjutnya data *insample* dan *outsample* ARIMA (1,0,2)(1,0,2)¹² digabungkan sehingga diperoleh nilai forecast, kemudian dikembalikan ke bentuk awalnya sebelum ditransformasi. Nilai ini digunakan sebagai nilai peramalan 6 bulan mendatang, yaitu bulan Juli 2010 sampai dengan Desember 2010 (Tabel 4).

PEMBAHASAN

Metode *time series* merupakan metode peramalan yang melalui analisis suatu variabel yang diperkirakan dengan variabel atau fungsi waktu. Langkah penting dalam pemilihan



Gambar 1. Mean dan Standar Deviasi Keseluruhan Kasus



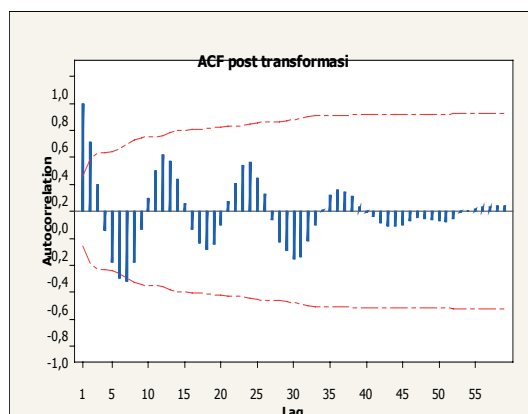
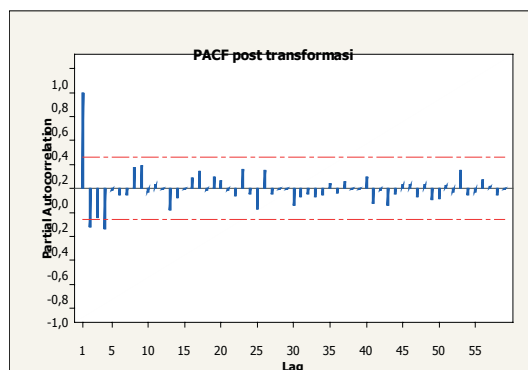
Gambar 2. Plot Time Series setelah Transformasi

Tabel 3. MAPE, MAD, MSD Data *Insample* dan *Outsample* Metode *Winters*

	<i>Insample</i>				<i>Outsample</i>			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
MAPE	49,239	49,1487	49,14212	49,85649	3,9906	3,9946	4,088	3,8810
MAD	92,5707	92,5058	88,5205	102,0127	18,0822	18,1006	18,5411	17,4669
MSD	20012,5	19993,2	18322,02	24936,64	4822,085	4831,233	5041,653	4535,98

Tabel 4. MAPE, MAD, MSD Data *Insample* dan *Outsample* Metode ARIMA

	<i>Insample</i>			<i>Outsample</i>		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
MAPE	3,9667	5,0891	4,056799	1,2237	1,0286	1,3131
MAD	0,1935	0,2523	0,200005	0,0725	0,0620	0,0776
MSD	0,067899	0,092537	0,06678	0,058487721	0,0489032	0,066328722

**Gambar 3.** ACF kasus DBD di Kota Surabaya Januari 2005–Desember 2009**Gambar 4.** PACF kasus DBD di Kota Surabaya Januari 2005–Desember 2009

metode peramalan yang tepat adalah dengan mempertimbangkan jenis pola data. Dalam penelitian ini dibahas dua metode peramalan yaitu metode *Winters* dan metode ARIMA. Menurut Box dan Jenkins (1976), metode ARIMA (*Autoregressive Integrated Moving Average*) atau juga dikenal dengan metode Box-Jenkins merupakan salah satu bentuk

analisis deret waktu (*time series analysis*). Model ARIMA dapat menganalisis data secara univariat yang mengandung pola musiman maupun *trend*. Metode ini hanya menganalisis data yang stasioner, sehingga data yang tidak stasioner harus distasionerkan terlebih dahulu dengan transformasi atau pembedaan. Sedangkan pada metode pemulusan eksponensial yang

Tabel 5. Hasil Ramalan (Aktual) 6 Bulan Mendatang Metode ARIMA

Periode	FITS (T)	FORE (T)	FITS (A)	FORE (A)
1	4,26804	5,65264	22,7705828	46,49607015
2	4,34350	5,14426	23,73556703	36,17454923
3	5,39772	4,67609	41,05849257	28,39196078
4	5,01943	4,37538	33,94854911	24,15307454
5	4,78294	3,54382	30,03563914	15,02240657
6	5,27031	4,06096	38,54081825	20,2845341

dapat menganalisis berbagai data baik data yang stasioner maupun data yang tidak stasioner dengan pemisahan aspek *trend* dan musiman (Kurniawan, 2008).

Langkah pertama yang baik untuk menganalisis data runtun waktu adalah memplot data tersebut secara grafis. Hal ini bermanfaat untuk menetapkan adanya *trend* (penyimpangan nilai tengah) untuk mengetahui adanya pengaruh musiman pada data (*deseasonalize the data*) (Makridakis et al., 1995).

Metode peramalan ARIMA dan *Winter* dapat mengatasi persoalan data yang mengikuti pola musiman dan trend. Metode *Winter* lebih sederhana dibandingkan dengan metode ARIMA baik dalam tahapan maupun dalam menentukan parameter yang harus dimasukkan ke dalam rumusan metode.

Pada deskripsi statistik dari jumlah kasus DBD terlihat bahwa rata-rata jumlah kasus DBD paling tinggi dicapai pada bulan April. Pada bulan April kota Surabaya sedang berada pada musim penghujan. Variansi yang terbesar terjadi pada bulan Maret.

Hal ini berarti jumlah penderita DBD pada bulan ini sangat bervariasi, pernah terjadi kasus mencapai angka yang tinggi dan pernah pula pada bulan ini angka kejadian kasus yang rendah, meskipun tidak serendah pada bulan-bulan yang mana rerata angka kejadian kasusnya rendah. Jumlah kasus DBD tercatat tertinggi pada tahun 2006 terjadi pada Maret, yaitu sebesar 841 kasus, sedangkan jumlah kasus terkecil yakni sebesar 40 kasus terjadi pada bulan Oktober.

Dalam proses dan strategi pembentukan model, pengumpulan data merupakan hal

merupakan hal yang sangat penting dan sangat menunjang di dalam hal menganalisis data. Data yang dikumpulkan dibagi menjadi 2 yaitu *in-sample* dan *out-sample*. Data pertama (*in-sample*) untuk pembentukan model yang terdiri dari data bulanan jumlah kasus DBD Januari 2005–Desember 2009 (60 bulan). Data kedua (*out-sample*) untuk evaluasi peramalan yang terdiri dari data bulanan jumlah DBD bulan Januari–Juni 2010 (6 bulan).

Identifikasi Model dengan Metode *Winter*

Metode *Winters* merupakan salah satu varian metode pemulusan. Dasar dari metode pemulusan adalah pembobotan atau pemulusan pengamatan masa lalu dalam suatu deret berkala, untuk memperoleh ramalan masa mendatang.

Metode *Winters* yang selama ini dikenal dibedakan atas dua, yaitu *Winters* aditif dan *Winters* multiplikatif⁸. Perbedaan antara metode *Winters* aditif dan multiplikatif terdapat pada asumsi bagaimana komponen musiman mempengaruhi deret waktu. Metode *Winters* aditif biasanya cocok untuk suatu deret waktu yang amplitudo pengaruh musimannya tidak tergantung pada tingkat mean suatu deret. Metode *Winters* multiplikatif biasanya digunakan untuk data time series yang amplitudo pola musimannya sesuai dengan tingkatan deretnya (Terasvirta, 1993).

Dalam penelitian ini digunakan metode *Winter* multiplikatif karena dari plot time series diperoleh Gambaran yang bersifat multiplikatif. Adapun metode *Winter* multiplikatif yang digunakan mempunyai parameter α , β , dan γ . Parameter tersebut kemudian ditentukan dengan

cara coba-coba sehingga diperoleh *galat error* (MAPE, MAD dan MSE) yang terkecil di antara keempat model yang dicoba. Oleh karena metode ini mempunyai kendala yaitu membutuhkan banyak waktu maka metode ini jarang digunakan. Setelah memperoleh nilai galat yang terkecil maka ditentukan model terbaiknya dan selanjutnya dengan menggabungkan seluruh data lampau dan data aktual diperoleh nilai peramalan untuk 6 bulan mendatang. Dalam hal ini dipakai waktu 6 bulan karena makin pendek waktu yang diramalkan maka keakuratan hasilnya semakin tinggi.

Identifikasi Model dengan Metode ARIMA

Sebelum data dianalisis dengan menggunakan metode ARIMA, data time series terlebih dahulu diidentifikasi apakah data sudah stasioner di dalam mean dan varians. Dari hasil analisis langkah dan strategi pembentukan model dengan metode analisis deret berkala ARIMA nampak bahwa plot data asli jumlah kasus demam berdarah di Surabaya sudah stasioner dalam mean namun tidak stasioner dalam varian. Karena data belum stasioner dalam varians, maka harus dilakukan transformasi untuk menghilangkan ketidakstasionerannya tersebut (Makridakis, 1999).

Adapun cara untuk menguji stasioneritas suatu deret berkala adalah dengan menghitung nilai autokorelasi dari deret data asli. Apabila nilai tersebut turun dengan cepat ke atau mendekati nol sesudah nilai kedua atau ketiga menandakan bahwa data stasioner di dalam bentuk aslinya. Sebaliknya, apabila nilai autokorelasinya tidak turun ke nol dan tetap positif menandakan data tidak stasioner sehingga diperlukan pembedaan atau transformasi terhadap data asli. Jadi suatu data dinyatakan stasioner bila memenuhi persyaratan sebagai berikut: 1) apabila suatu data berkala diplot dan kemudian tidak terbukti ada perubahan nilai tengah dari waktu ke waktu, maka dikatakan bahwa deret tersebut stasioner pada nilai tengahnya (*mean*), 2) apabila plot data berkala tidak memperlihatkan ada perubahan varians yang jelas dari waktu ke waktu, maka dapat dikatakan deret data tersebut stasioner pada variansnya.

Apabila kestasioneran telah diperoleh, nilai autokorelasi dihitung untuk mengetahui pola lain

yang terkadang terdapat di dalamnya (yaitu selain keacakan yang tersebar di sekitar nol). Terdapat 3 kemungkinan yang harus dipertimbangkan, yaitu: 1) mungkin faktor musiman tampak dengan sendirinya. Nilai auto-korelasi untuk time-lag setiap kuartal atau setiap tahun besar dan secara signifikan berbeda dari nol, 2) mungkin terungkap adanya proses AR atau MA. Pola dari autokorelasi dan autokorelasi parsialnya akan menunjukkan suatu model yang memungkinkan, 3) campuran dari ketidakstasioneran, musiman dan proses AR dan MA mungkin akan terlihat (model ARIMA yang umum).

Data *insample* maupun *outsample* yang belum stasioner dalam varians tersebut selanjutnya ditransformasi menggunakan transformasi Box Cox. Transformasi *Box Cox* adalah transformasi pangkat pada respon. *Box Cox* mempertimbangkan kelas transformasi berparameter tunggal, yaitu λ yang dipangkatkan pada variabel respon Y^{λ} , sehingga transformasinya menjadi Y^{λ} , adalah parameter yang perlu diduga. Dengan transformasi Box Cox secara simultan dapat dicapai kenormalan dari sebaran, kekonstanan ragam dari galat dan linieritas dari struktur model. Bila dalam menduga parameter λ kurang tepat, maka akan dicirikan oleh lebarnya selang kepercayaan.

Setelah dilakukan transformasi terhadap data asli, maka langkah selanjutnya adalah melihat nilai ACF dan PACF dalam rangka penstasioneran data (Anugrah, 2007). Dari Gambar 2 terlihat bahwa plot data tersebut sudah stasioner dalam mean maupun varian. Hasil visual juga menunjukkan bahwa varians data terkoreksi menjadi lebih baik (perbedaan tinggi rendah puncak satu dengan puncak lainnya semakin kecil). Hal tersebut menunjukkan hasil transformasi tersebut berhasil. Namun untuk lebih memastikan keberhasilan transformasi *Box Cox* dalam memperbaiki data agar stasioner dalam varians maka dilakukan uji kesamaan varians ANOVA. Uji ini berguna untuk menentukan apakah data kasus demam berdarah dengue tersebut homogen atau tidak dalam varians.

Data transformasi selanjutnya diimplementasikan dengan program komputer sehingga diperoleh plot ACF dan PACF. Dari bentuk ACF dan PACF tersebut dapat diestimasi beberapa model ARIMA yaitu (2,0,2)(2,0,2);

(1,0,2)(1,0,2) dan (1,0,1)(1,0,1) dengan periode 12 bulan.

Setelah berhasil menetapkan identifikasi model sementara, selanjutnya parameter-parameter AR dan MA, musiman dan tidak musiman harus ditetapkan. Uji signifikansi parameter dilakukan dengan statistik uji t , sedangkan uji apakah error sudah white noise dilakukan dengan statistik uji *Ljung-Box Chi square*.

Langkah selanjutnya adalah diagnosis model. Diagnosis model dilakukan untuk mendeteksi adanya korelasi dan kenormalan antar *residual*. Dalam runtun waktu (*time series*) ada asumsi bahwa *residual* mengikuti proses *white noise* yang berarti *residual* harus independen (tidak berkorelasi) dan berdistribusi normal dengan rata-rata mendekati 0 ($\mu = 0$) dan standar deviasi (σ) tertentu (Iriawan, 2006: 361). Untuk mendeteksi adanya proses *white noise*, maka perlu dilakukan diagnosis model dengan cara uji independensi *residual*. Uji dilakukan untuk mendeteksi independensi *residual* antar lag. Dua lag dikatakan independen (tidak berkorelasi) apabila antar lag tidak ada korelasi cukup berarti.

Dalam penelitian ini, uji dilakukan dengan membandingkan 2 *Ljung-Box* χ^2 dan nilai Tabel $\chi^2_{(\alpha, df)}$ pada output proses *Ljung-Box-Pierce*.

Hipotesis: H_0 : $\rho_{at+at+k} = 0$ (Ada korelasi antar-lag) dan H_1 : $\rho_{at+at+k} \neq 0$ (Tidak ada korelasi antar lag atau minimal ada 1 lag yang $\neq \rho_{at+at+k}$)

Kriteria penolakan H_0 yaitu jika 2 *Ljung-Box* $\chi^2 < \chi^2_{(\alpha, df)}$, di mana distribusi χ^2 yang digunakan mempunyai $df = k-2$.

Selain dengan pengujian hipotesis, independensi antar lag akan ditunjukkan pula oleh grafik fungsi autokorelasi (fak) *rasidual*. Suatu *residual* model dikatakan telah independen jika tidak ada satu lag pun pada grafik fungsi autokorelasi (fak) *rasidual* yang keluar batas garis (Ispriyanti, 2004).

Langkah selanjutnya adalah melakukan estimasi parameter dari model ARIMA dugaan ini, yang selanjutnya masing-masing dilakukan evaluasi untuk menentukan model terbaik diantara dugaan yang ada, yaitu ARIMA (2,0,2)(2,0,2), ARIMA (1,0,2)(1,0,2) dan ARIMA (1,0,1)(1,0,1). Berdasarkan Tabel 3 diperlihatkan

bahwa dari ketiga model *insample* yang diperoleh model yang terbaik adalah model 1, yaitu ARIMA (2,0,2)(2,0,2). Model 1 memiliki MAPE, MAD dan MSD yang terkecil dibandingkan dengan kedua model data *insample* lainnya. Dari ketiga model data *outsample* diperoleh hasil bahwa model yang terbaik adalah model 2, yaitu ARIMA (1,0,2)(1,0,2) karena mempunyai MAPE, MAD dan MSD yang terkecil bila dibandingkan dengan kedua model data *outsample* lainnya. Dalam kasus ini karena ada perbedaan nilai terbaik antara data *insample* dan data *outsample*, maka yang dipake sebagai model terbaik adalah model 2, karena data *outsample* adalah data yang lebih baru, sehingga dianggap lebih memproyeksikan kondisi yang akan diramalkan. Di samping itu model 2 juga mempunyai orde yang lebih kecil bila dibandingkan dengan model data *insample* maupun *outsample* ARIMA (2,0,1)(2,0,0) selanjutnya digabungkan sehingga diperoleh nilai forecast yang digunakan sebagai nilai peramalan 6 bulan mendatang, yaitu untuk Juli 2010 sampai dengan Desember 2010 (Tabel 4).

Dari keseluruhan hasil penelitian didapatkan bahwa Metode ARIMA (1,0,2)(1,0,2) adalah metode terbaik untuk menggambarkan dan meramalkan pola atau fluktuasi jumlah kasus DBD dibandingkan dengan metode *Winters*. Hal ini disebabkan karena metode *Winters* menggunakan parameter yang ditentukan secara coba-coba (White, 1989), tidak ada dasar tolok ukurnya, sedangkan metode ARIMA ada kriteria yang dapat ditentukan dengan ACF dan PACF. Hasil ini juga memberikan suatu fakta bahwa metode yang lebih kompleks (rumit), dalam hal ini ARIMA yang terdiri atas tahap identifikasi, estimasi dan forecasting, cenderung memberikan hasil yang lebih baik pula.

Model peramalan metode *Winters* dengan parameter α (level) = 0,2; γ (trend) = 0,15; dan δ (seasonal) = 0,002 adalah

Pemulusan Eksponensia $I_t = 1$

$S_t = 0,2(X_t/I_t) + 0,8(S_{t-1} + b_t)$

Pemulusan Trend

$b_t = 0,15(S_t - S_{t-1}) + 0,85b_{t-1}$

Pemulusan Musiman

$I_t = 0,002(X_t/S_t) + 0,998I_{t-1}$

Peramalan

$F_{t-m} = S_t + b_t.m)I_{t-L+m}$

Sedangkan model peramalan Metode ARIMA (1,0,2)(1,0,2).

Dari hasil yang diperoleh diketahui bahwa peningkatan jumlah kasus DBD yang hampir selalu terjadi pada bulan yang sama setiap tahunnya, kecuali pada bulan Agustus tahun 2005. Hal ini perlu mendapat perhatian dari pemerintah sebagai penyanggah dana maupun tenaga kesehatan sebagai garda terdepan yang berhadapan langsung dengan kasus. Selain itu harus didukung oleh sarana dan prasarana kesehatan yang memadai. Kasus sempat mencapai puncaknya pada bulan Maret tahun 2006 sebelum akhirnya mengalami penurunan pada tahun berikutnya. Pola musiman memberikan Gambaran bahwa musim penghujan dan masa pergantian musim penghujan-kemarau (musim pancaroba) secara dominan sangat berpengaruh terhadap peningkatan kejadian DBD di Surabaya.

Hasil ini dapat digunakan untuk membantu pihak-pihak yang terkait dalam penanganan dan penanggulangan penyakit DBD yang berkaitan dengan waktu sehingga dapat dilakukan pencegahan maupun penanganan yang optimal. Pola ini juga dapat digunakan oleh pihak Puskesmas dan rumah sakit untuk mempersiapkan sarana dan prasarana guna meminimalkan angka morbiditas dan mortalitas.

SIMPULAN

Setelah melakukan peramalan dengan metode Pemulusan (Smoothing) *Winters* dan ARIMA maka dapat disimpulkan bahwa Metode ARIMA (1,0,2)(1,0,2) adalah metode terbaik untuk menggambarkan dan meramalkan pola atau fluktuasi jumlah kasus DBD dibandingkan dengan metode *Winters*.

DAFTAR PUSTAKA

- Box, G.E.P., & Jenkins, G.M. 1976. *Time Series Analysis, Forecasting and Control*. Holden-Day: San Francisco.
- Depkes RI. 2006. *Pemberantasan Sarang Nyamuk Demam Berdarah Dengue (PSN DBD) oleh Juru Pemantau Jentik (Jumantik)*. Ditjen PP & PL: Jakarta.
- Depkes RI. 2005. *Pencegahan dan Pemberantasan Demam Berdarah Dengue di Indonesia*. Ditjen PP & PL: Jakarta.
- Depkes RI. 2004. *Kajian Masalah Kesehatan Demam Berdarah Dengue*, Jakarta. www.litbang.depkes.go.id, (sitasi 11 Desember 2009).
- Ispriyanti, D. 2004. Pemodelan Statistika dengan Transformasi Box Cox, *Jurnal Matematika dan Komputer*, Vol. 7. No. 3, 8–17.
- Kurniawan, D. 2008. *Regresi Linier*, <http://ineddeni.wordpress.com>. (sitasi 1 Februari 2010).
- Makridakis, S., Wheelwright, S. & Victor, E.M. 1999. *Metode dan Aplikasi Peramalan*, 2nd edition, Erlangga: Jakarta.
- Makridakis, S., Syeven C.W & Victor, E.M. 1995. *Metode dan Aplikasi Peramalan*, Terjemahan Hari Suminto, Binarupa Aksara: Jakarta.
- Terasvirta, T., Lin, C.F & Granger, C.W.J. 1993. Power of The Neural Network Linearity Test, *Journal of Time Series Analysis*, Vol. 14, p. 159–171.
- White, H. 1989. An additional Hidden Unit Test for Neglected Nonlinearity in Multilayer Feed forward Networks. In *Proceedings of The International Joint Conference on Neural Networks*, pp. 451-455, San Diego, CA: SOS Printing, Washington, DC.